2016-03

Distance to Promotion: Evidence from Military Graduate Education

Bacolod, Marigee

Distance to Promotion: Evidence from Military Graduate Education*

Marigee Bacolod† Latika Chaudhary‡

First Draft: March 2016

Abstract

While a large literature has studied the effects of different education delivery modes on course grades and related education outcomes, we know much less about how these delivery modes impact subsequent job performance. Using a unique dataset of US military officers pursuing graduate education at the Naval Postgraduate School, we implement propensity score matching methods to identify the effects of distance education programs on academic and job outcomes such as promotion within the military and separation from the military. The distance education degree programs in our context share many common features with online programs that are growing in number by the day. We find a large and negative impact of attending distance education degree programs on average GPA, graduation, number of thesis extensions, promotion and separation from the military. That said, the average effects of distance education mask substantial heterogeneity by group. The negative effects are smaller among pilots and Hispanics compared to among junior military officers and individuals who earned their undergraduate degrees from public universities.

Keywords: Distance Education, Online Education, Returns to Education, Matching.

JEL Codes: I20, I23

*We thank NPS half-baked seminar participants for comments. All errors are our own. The views expressed in this article are those of the authors and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

†Associate Professor, Naval Postgraduate School, Email: mbacolod@nps.edu

‡Associate Professor, Naval Postgraduate School, Email: lhartman@nps.edu

1 Introduction

Online and distance education (DL) has exploded over the last decade in the United States and globally.[1] According to recent estimates, 11% of degree-seeking undergraduates are enrolled in complete online courses. While for-profit colleges have led the charge, public and nonprofit colleges are slowly catching up. Many of the latter had occasional online courses, but are now offering complete online degree programs similar to for-profit chains such as University of Phoenix (McPherson and Bacow, 2015). Students from around the world can now enroll in massive open online courses (MOOCs) offering the same course materials used by top US universities.[2] Given the lower cost of delivering online education, the transition away from traditional face-to-face instruction has the potential to lower education costs that have steadily increased over the past few decades (Deming, Goldin, Katz and Yuchtman, 2015).

With this growth in new technologies for delivering education, policymakers and academics continue to debate whether and how the medium of instruction influences education outcomes. Despite the large number of studies, there is no robust consensus that student learning in any one type of delivery mode clearly dominates the rest. A recent meta-analysis by the US Department of Education (Means et al. 2010) reviewed over 1,000 studies comparing student performance in online versus traditional settings. They concluded that of the 1,000, only 50 used credible randomization or quasi-randomization strategies to address the differential selection of students into online vs traditional courses. Within this group of studies student outcomes were higher in online courses compared to traditional courses, but the effects were economically small. Another meta analysis on just distance education (Bernard et al. 2004) found no significant differences in student outcomes between distance and traditional courses. But, the average masked heterogeneity in asynchronous and synchronous courses. Students enrolled in distance synchronous education performed worse than traditional students, while students enrolled in distance asynchronous courses performed better than traditional students.

Against this backdrop our paper uses a rich student-level dataset to study the impact of distance and online education on student educational outcomes and subsequent job performance. We follow multiple cohorts of military students who began graduate programs at the Naval Postgraduate School (NPS) between 2006 and 2013. NPS offers graduate programs in Business Administration, Engineering, Information Systems and other technical fields with a

[1] Hereafter, we refer to online and distance education as “distance learning” or DL.

[2] The non-profit company edX led by Harvard and MIT offers many top quality online courses to students worldwide.

combined distance and online option for many residence programs.[3] We match information on course grades, degree programs and graduation for each US military student with their demographic and career information reported in the database of the Defense Manpower Data Center (DMDC). Thus, we can track the impact of distance/online education on subsequent job outcomes such as promotion in the military, time to promotion, and attrition from the military.

Students enrolled in the distance education programs are located in military bases across the country. Their courses are either offered via synchronous video teleconferencing with recording for future views, or via web-based portals such as Elluminate. Both modes involve a live instructor in a separate location at the other end; the key difference is whether students are interacting with the instructor through video or online. We are currently collecting data to code each individual course into distance video teleconferencing and distance online. A future version of this draft will then be able to analyze the extent to which various modes of delivery of distance education matters. In this version, we restrict our focus to a simple indicator for students enrolled in distance education programs. Since students enroll in either the residence or distance learning option for the *entire* degree program, we focus on overall GPA as the main academic outcome in this version. Our future draft will be able to explore the impact on individual courses.

Our setting offers three key advantages. First, we track a large number of students (n=6,754). This is significantly more than studies that compare outcomes in single economics or statistics courses. Small sample sizes are also a concern in the few credible randomization studies because of low participation rates. Second, we link the mode of delivery to average GPA, graduation and subsequent job performance. This allows us to test for heterogeneous differences across degree programs and to study the impact of delivery mode on subsequent labor market outcomes. Following students beyond graduation is very rare in the literature. Third and finally, we use detailed background information to created matched samples of students to mitigate selection problems. The institutional bureaucracy of the U.S. Military also helps in the matching exercise because students are assigned to DL versus resident degree programs partially based on military needs that are not always correlated with individual student characteristics.

On the other hand, our setting is relatively unique in that the students in our analysis sample are all active duty officers across services of the U.S. Armed Forces. Both DL and resident students have their tuition paid for by their employers and draw a full-time salary while in school, unlike most civilian graduate students. Nonetheless, our setting offers a

[3] In 2013 NPS offered 81 Master's and 16 Ph.D. degree programs.

unique opportunity to derive causal estimates of DL's program impact not just on education outcomes but also on job success, which in turn have important implications for the growing cost of higher education.

Individual selection into DL or residence programs is the main challenge to identifying the impact of delivery mode on outcomes. If students are selecting into degree programs based on their ability for example, then a failure to control for ability will lead to biased estimates of distance education. We address selection using matching estimation techniques. The idea underlying matching is to use observable characteristics to match DL students to similar residence students in comparing their outcomes. In the first stage of matching, we predict enrollment into DL programs based on a rich set of individual characteristics including age, race, gender, marital status, service, military occupation category, military rank, type of undergraduate institution, and a crude proxy for ability in the form of an academic profile code assigned by the NPS admission office to each student applicant. Next we construct predicted propensity scores from the first stage regression for both types of students. In the second stage we estimate the impact of distance education on outcomes matching on these propensity scores. This ensures a more apples-to-apples kind of comparison in estimating the effect of DL. Our preferred estimation uses nearest neighbor matching with Abadie-Imbens standard errors that correct for using estimates of propensity scores in the second stage regression.

We find a large, negative and significant impact of distance education on average GPA using these matching methods. The average GPA of students enrolled in distance programs is 0.38 points lower than observationally similar resident students, which translates into a third of a letter grade. We also find that OLS regressions substantially underestimate the impact of distance education. In OLS specifications the negative impact of DL on average GPA is 0.27 points suggesting the OLS estimates are biased down.

Apart from GPA, we also find a negative impact of distance education on graduation and a positive impact on graduating late conditional on graduation. Distance education students are 15 percentage points less likely to graduate and 8 percentage points more likely to graduate late conditional on graduation. They are also likely to request more thesis extensions. In subsequent job performance, distance education students are 12 percentage points less likely to be promoted and are 9 percentage points more likely to separate from the military. These results are surprising in the context of a hierarchical organization such as the US military where explicit factors (for example, time in service) often guide promotion decisions.

When we test for heterogeneous differences, we find the negative effects of distance education vary significantly by group. We do not observe robust negative effects of distance programs for women, and also find smaller negative effects on Hispanics. Students who earned their undergraduate degree from a public institution do worse in distance programs compared to those from private non-profit institutions. Among military communities, pilots in distance programs do only marginally worse than resident pilots, but submarine officers do significantly worse in distance programs. These findings are broadly consistent with other settings where the average impact of mode of delivery exhibits tremendous heterogeneity.

The rest of the paper is organized as follows. We review the related literature in the next section. Section 3 describes our data and institutional setting. Section 4 discusses the matching estimating. We present the results in Section 5 and conclude in Section 6.

2 Related Literature

A large and growing literature has studied the impact of course delivery on learning outcomes. Older studies focused on distance education when universities had remote campuses for students who were unable or unwilling to enroll in traditional campuses with face-to-face instruction. As the technology of delivering course and degree content has evolved, several studies have compared the impact of pure online instruction, blended instruction (combining some elements of online instruction with some face-to-face meetings), and traditional instruction on student outcomes as captured by either test scores or course grades.

There are several meta-analyses of how the mode of delivery impacts student outcomes. Since the NPS distance learning degree programs incorporate elements of both distance and online education, we begin by discussing two relevant meta-analyses. An important study by Bernard et al. (2004) compares the performance of distance education to traditional instruction. A defining characteristic of distance education is instructors being in separate locations from their students using some form of video teleconferencing for either lecture delivery or office hours. That said, in recent years distance education has embraced aspects of online education with taped lectures posted on course websites, online course assignments and tests.[4] So distance education today can be viewed as a type of online education.

Distance education courses are either offered as synchronous courses with the instructor teaching in the same fashion as in a traditional course via a video teleconferencing program and students viewing the lecture at the same time. In contrast, asynchronous distance education courses allow students to watch previously recorded lectures at their own convenience.

[4]For example, the platform Aplia offers online assignments for many economics courses that are used at NPS by instructors offering online, traditional and distance education courses.

Although Bernard et. al (2004) find no significant difference in outcomes between distance and traditional courses, the average effects are misleading because of differences between asynchronous and synchronous courses. Asynchronous distance education is correlated with better student outcomes compared to traditional courses, while synchronous distance education leads to worse student outcomes. Their study also notes the wide range of estimates in the literature that are often driven by different estimation strategies. The older studies tend to use observational data without statistically correcting for the selection of students into courses, making it difficult to draw strong conclusions.

A recent meta-analysis by the US Department of Education (2010) comparing online and traditional instruction explicitly focuses on a small number of studies with a randomized or quasi-randomized study design. They find no significant difference in outcomes between online and traditional courses, but they find small and positive effects of blended courses on outcomes relative to traditional courses.

Randomized-control trials remain the gold standard for estimating causal effects of policies, programs, or in this case the mode of delivery. Although the number of randomized studies is increasing (Bowen, Chingos, Lack and Nygren 2014; Joyce, Crockett, Jaegar, Altindag, and O'Connell 2014; Figlio, Rush and Yin, 2014; and Alper, Couch and Harman, 2015), randomized experiments in this particular context have their shortcomings. First, the sample sizes tend to be small because of low participation rates. Recent efforts by Figlio et al. (2014) get around this problem by offering students extra credit to participate in the experiment. On a related note, external validity is also a problem because most randomization studies focus on introductory economics or statistics courses. Second, these experiments suffer from contamination because it is quite likely that freshmen enrolled in the same section, for example, of an introductory economics course speak to each other about their class experiences (online or traditional). Survey questions rarely ask about such interactions. Third, students in traditional courses access all sorts of online materials so these studies are rarely estimating the impact of pure online instruction because we do not know how many students access online materials (for example, Khan Academy videos on YouTube). Despite these limitations, the randomized studies tend to find no significant differences on average between the different instructional modes. However, these averages mask important heterogeneity. Students from disadvantaged backgrounds do worse in online and blended classrooms compared to under traditional instruction (Alper, Couch and Harman, 2015).

A smaller set of studies within economics has used observational data with corrections for the selection problem. For example Coates, Humphreys, Kane and Vachris (2004) had students take tests before and after enrolling in an introductory principles course across three campuses. Students chose whether to take the course online or face-to-face, but the authors used a rich background survey instrumenting for course choice using commute time to campus and general familiarity with online platforms. Correcting for selection, they found students perform worse in online courses. Importantly, they show that their OLS estimates suggest no significant difference in outcomes, which highlights the importance of correcting for selection when using observational data.

In a similar vein Olitsky and Cosgrove (2015) combine difference-in-difference with matching to address the selection problem. Using test scores at multiple points in time, they also find no significant differences in outcomes between blended and traditional courses. However, they find that students in traditional introductory economics courses have higher content retention compared to students enrolled in blended classrooms.

Our study is both similar and different to these observational studies. We also use rich demographic student information combined with propensity score matching to correct for selection into degree type like Olitzky and Cosgrove (2015). Unlike these other studies, however, we study several student outcomes from course grades and graduation to labor market outcomes. Since our students enroll in a distance education or resident degree program, we cannot separately identify the impact of mode of delivery on individual course grades. This is especially true for courses in a sequence where the online mode may work well for one course in the sequence but not the other. We show the range of impacts on individual course grades for completeness, but our analysis focuses on the overall GPA, time to graduation and graduation. We are also unique in studying labor market outcomes – promotion within the military, time to promotion, and attrition from the military. A recent study using fictitious resumes finds that online business degree holders are less likely to get job call backs (Deming et al. 2016). Although very interesting, their study speaks more about the perception of online degrees in labor markets. In contrast, our study focuses on online degrees that are offered by an accredited university that offers online and resident programs to military students. We speak directly to the impact of such non-resident degrees on subsequent labor market success.

3 Institutional Background and Data

NPS is a unique institution offering employer-funded graduate education tailored to support facets of US national defense. Majority of the student body are military officers from all branches of the U.S. uniformed services. The rest of the student body are civilian employees of the federal government, and military officers and civilians from about 47 U.S. allied countries.

The selection of U.S. military officers for graduate studies at NPS varies by branch of service, but is usually based on outstanding job performance, promotion potential, academic background, and military needs. NPS determines academic eligibility, while each service has a selection board or an assignment officer that screens eligible officers for selection to NPS. The service-specific board assigns students to the degree program designed to prepare them for future career assignments, and based on military needs, assigns them to the resident or distance version of that degree program. Upon completing their degree, NPS students are required to fulfill an obligated service commitment with their branch of service.

An important consideration is that distance learning students are actively engaged in their job tasks when they are not in class. While resident students at NPS are still required to perform military-specific tasks such as muster, stand on duty, and train in physical fitness, resident students have fewer job tasks compared to DL students. In our matching framework below, we attempt to control for this differential workload by matching on officer rank and military community/occupation groups. While this difference is explicit in our context, DL students in non-miliary contexts are more likely to be employed while enrolled compared to resident graduate students.

Our data are drawn from the population of resident or traditional and non-resident or distance learning Master's degree-seeking students at NPS who began their graduate programs during Academic Years 2006 through 2013. Student records at NPS do not, however, include basic demographic information that have been shown in the literature to reliably predict students' educational outcomes (e.g. race and gender). Thus, we turn to the Personnel Master files housed at the Defense Manpower Data Center (DMDC). Because demographic and job performance information on civilian and international students are not reported in DMDC, our analysis excludes NPS students who are civilians or are international students (approximately 38% of our sample) leaving us with just under 7,000 students.[5]

Our demographic information such as marital status and number of dependents are drawn as of six months prior to students beginning their studies at NPS. In addition, we track our students' subsequent job performance upon departure from NPS, and collect information on their dates of promotion and/or separation, also from DMDC.

To capture characteristics of a student's prior education, we merge in information from the National Center of Education Statistics' Integrated Postsecondary Education Data System

[5]International students are exclusively enrolled in the residence programs, while civilians are disproportionately represented in the distance programs. Civilians accout for 63% of students in distance programs and 10% of students in residence programs.

(IPEDS) based on that military officer's undergraduate institution. The main variable that captures student preparation prior to arrival at NPS, however, is the student's Academic Profile Code (APC). The APC code is a three-digit code assigned by NPS' Admissions office when it evaluates applicants. The first digit is a categorical indicator of a student's academic performance based on their cumulative grade point average on all previous college transcripts. The second digit indicates the student's mathematics background, ranging from graduation as a Math major in the last 7 years to no pertinent college-level math. The third APC digit indicates a student's science/engineering/technical background. Each curriculum at NPS specifies their APC threshold for admission into their program. Appendix Table 1 provides a more extensive description of these indicators.

Table 1 reports summary statistics of our analysis sample. As is clear from the table, DL students, on average, have worse academic outcomes than resident students. DL student GPAs are 0.3 points below resident students, a difference that corresponds to one-third of a letter grade. DL students are 17 percentage points less likely to graduate. Conditional on graduation, they are also more likely to request thesis extensions. After NPS, DL students are 14 percentage points less likely to get promoted than resident students and are 2 percentage points more likely to separate from military service.

In terms of demographics, DL students are different from resident students in a few ways. Women and minorities are relatively more represented in resident programs, at 9.6% female, 6.6% Black and 6.9% Hispanic, than in distance education where they constitute only 5.3%, 4.8% and 6.2%, respectively. Marital status and number of dependents are similar across the two groups. Distance students are a tad younger on average, but there is also more variation in their age distribution than among residents.

In terms of job-related characteristics, Navy students are disproportionately represented in the distance programs at 89% compared to 51% in residence programs. Within occupational communities, pilots are more likely to enroll in distance education often so they can maintain flight hours, a necessary requirement for their jobs. In contrast, intelligence and support officers are more likely to attend residence programs. A majority of students in both programs are early career officers of rank O3, which corresponds to a Lieutenant in the Navy and a Captain in the other services.

Finally, Table 1 shows DL students are on average less academically prepared than resident students. There are significantly more DL students who did not meet the academic requirements of their degree programs. For instance, only 63.5% of DL students met the mathematics requirement of their curriculum compared to 76.8% of resident students. APC waivers are typically granted on a case-by-case basis by program administrators at NPS. That said, military students from service academies that generally would be considered selective undergraduate institutions are more likely to be represented among distance students.

The broad differences between the two groups do not provide evidence of strong positive or negative selection into either degree mode. On the one hand, distance students are less likely to meet academic requirements of their programs, which would suggest negative selection into distance education. However, they are also drawn from service academies and the aviator community that would be evidence of positive selection. The service academies are selective institutions and military pilots often score high on ability tests. However the differences between the groups highlight the need to seriously address selection problems. We turn to this topic next.

4 Matching Estimation

The objective of our estimation strategy is to identify the causal impact of distance education on student educational outcomes and subsequent job performance. This is, however, fraught with empirical difficulties. It is important to control for each student's characteristics that influence both his/her assignment to a resident vs. DL program as well as his/her educational and subsequent job performance outcomes. For example, if students who are more motivated and productive tend to attend resident programs, then the negative correlation between DL and outcomes seen in Table 1 is attributable to this unobserved student characteristic. At the same time, positive selection towards distance learning may occur. This would be the case if the more productive students (e.g. those who are more able to manage their time and learn efficiently) tend to be in DL programs.

As noted in the introduction, an appeal of our setting is that assignment to a distance vs. resident degree program depends mainly on military needs rather than individual student characteristics. For instance, the Aviation community–comprised mostly of pilots who need certain flight hours to maintain certification on particular craft–is over-represented in our DL sample at 52% while constituting only 14% among our resident sample (see Table 1).

Our empirical strategy is designed to address these selection issues while controlling for student heterogeneity. Following the program evaluation literature, our goal corresponds to estimating the difference in potential outcomes, averaged across all students. That is, we would like to estimate the Average Treatment Effect (ATE) of DL:

$$\tau = E[Y_i(1) - Y_i(0)]$$

where $Y_i(1)$ is the potential outcome of student i if he/she was enrolled in a distance learning program while $Y_i(0)$ is his/her outcome if he/she was in a resident degree program. However, we can only observe

$$Y_i = \begin{cases} Y_i(0) & if\, DL_i = 0 \\ Y_i(1) & if\, DL_i = 1 \end{cases}$$

To estimate the ATE, we form counterfactuals using propensity score matching. Introduced by Rosenbaum and Rubin (1983), this involves estimating the propensity of student i to be assigned to a distance learning program

$$p_i(X) = Pr(DL_i = 1|X_i),$$

where X_i denotes student i's observable characteristics. Under the assumptions of unconfoundedness and common support, that is, (i) $Y(1), Y(0) \perp DL|X$ and (ii) $p_{min} \leq p(X) \leq p_{max}$, we can form matching estimators to τ where

$$\hat{Y}_i(0) = \begin{cases} Y_i & if\, DL_i = 0 \\ Y_j & if\, DL_i = 1 \end{cases}$$

and

$$\hat{Y}_i(1) = \begin{cases} Y_j & if\, DL_i = 0 \\ Y_i & if\, DL_i = 1 \end{cases}$$

and Y_j is the outcome of student j whose propensity score, or predicted probability of being assigned to a DL program, is the nearest to student i. Note that we match with replacement so the same units may be used as a match more than once. Abadie and Imbens (2006, 2010) show that the bias properties of the simple nearest matching estimator can be removed by a regression adjustment. The adjustment is necessary because the propensity score is an estimate in the first place. To implement this, we utilize STATA 13.0's `teffects psmatch` algorithm.

There are of course several ways of implementing the matching method. One can weight the observations by the estimated propensity score (also known as propensity score weighting). Alternatively, one can use the propensity score directly as a regressor, or divide the sample into subsamples with approximately the same values of the propensity score (also known as blocking). We implement various alternatives as robustness checks and also for comparison with our nearest matching estimator.

In terms of a parametric formulation, we estimate the linear regression

$$Y_i = \alpha + \beta' X_i + \tau DL_i + \varepsilon_i,$$

where Y_i are the outcomes: overall GPA, graduation, late graduation (if graduated), number of thesis extensions (if graduated), promoted, separated. Below, we report our estimates of τ under various implementations of the matching method: (1) OLS restricting the sample where the propensity scores strictly overlap, (2) propensity score weighting, (3) propensity score weighting restricting the sample where the propensity scores strictly overlap, and our preferred estimator, (4) propensity score matching with correct standard errors described above.

5 Results

5.1 Propensity Score Estimation

Before we present our main results, we first discuss key aspects of our matching analysis. Table 2 presents logit regressions with an indicator for distance education as the dependent variable against individual student characteristics. We report these estimates for the entire sample in column (1) and for the sample of graduates in column (2). Women are significantly less likely to enroll in distance programs with a stronger effect among the graduate sample, suggesting women are more likely to graduate from distance programs compared to men. Married individuals are significantly more likely to enroll in distance programs in both samples. Although the coefficient on age is insignificant, we believe age is correlated with years since an undergraduate degree and military rank. Both these variables are correlated with distance education - resident students earned their undergraduate degree more recently than distance students. Residents are also more likely to be junior officers (O2s), while distance students are more likely to be senior officers (O5s).

Given the substantial differences between DL and resident students shown in Table 1 and the estimates in Table 2, one may be concerned we do not have sufficient overlap in our data to implement matching. Overlap in the propensity scores across both groups, also known as having common support, is a key assumption of matching methods. Figures 1 and 2 assess the common support assumption using predicted propensities of students assigned to DL programs. We predict these propensities using the coefficient estimates from Table 2. An inspection of Figures 1 and 2 show a significant overlap in the propensity scores for both the overall sample and the sample of graduates. As expected, the marginal distribution for the control group (resident students) are massed near 0, while the distribution for the treatment group are shifted to the right. However, there is still substantial overlap over the interval [0,1]. If the two distributions did not overlap, then we would be unable to implement matching.

We also explore the extent of overlap within various sub-samples. These propensity scores are shown in Figures 3 to 7. As is clear from the pictures, we have sufficient overlap to test for the impact of distance education among females, whites, hispanics and married individuals. However, we seem to have less common support among blacks. We also seem to have overlap within different institution types (public, private non-profit, private for-profit and service academy) and by prior academic preparation. In terms of services, we have common support within the Navy and Marine Corps samples, but less so within the Army and Air Force. Finally, we have sufficient overlap within the aviation and submarine communities versus ground combat and intelligence where we observe few individuals in DL programs.

Overall, the idea underlying matching is to compare individuals within the area of common support who share similar propensities to be enrolled in distance programs. Our matching exercise uses a rich set of observables such as age, race, gender, military service, occupation and the first year of enrollment to generate propensity scores. Conditional on these observables, the assumption is that assignment to distance programs is uncorrelated with other factors that may impact academic performance and labor market outcomes.

5.2 Matching Estimates of Effect of Distance Education on Outcomes

Table 3 presents our results for the effect of DL on academic and job related outcomes. We report our results using multiple estimation techniques. Column (1) reports OLS estimates restricting the sample to where the propensity scores strictly overlap. We refer to this as the "Matched sample," where we exclude observations of resident students whose predicted probabilities of being in DL are strictly below the minimum of the propensity scores among DL students and observations of DL students whose scores are above the maximum among resident students. The matched sample effectively drops extreme outliers ($n = 65$), individuals that are very different from each other based on observables and their propensities for distance assignment.

Meanwhile, estimates in Column (2) use inverse of the propensity scores themselves as weights in the second stage regression over the whole sample.[6] In column (3), we also use the inverse of propensity scores as weights, but restrict our analysis to the matched sample used in column (1).

However, the inverse probability estimation, along with other matching methods, fail to

[6]More precisely, we weight DL students by $\frac{1}{\hat{p}_i(X)}$ and resident students by $\frac{1}{1-\hat{p}_i(X)}$.

take into account that the propensity score is itself estimated in a first step. Abadie and Imbens (2006, 2010) show this affects the large sample distribution of matching estimators and can generate biased standard errors. In column (4), we present matching estimates using nearest neighboring matching and Abadie-Imbens derived standard errors that account for the estimated propensity scores. Hence, our preferred estimates are those shown in column (4). Each cell in the table represents the coefficient on distance education from a regression estimation as shown in the columns for the dependent variable described in the rows. For example, the first cell with a coefficient of -0.2722 represents the coefficient on distance education with GPA as the dependent variable using OLS estimation on the matched sample.

We find that students enrolled in distance programs have a significantly lower GPA. The effects range from 0.3 points using OLS to 0.5 points using inverse probability matching in column (3). As we move across columns, the coefficient on DL increases suggesting that our OLS estimates are biased down. This suggests students are positively selected into the distance programs perhaps on account of their time management skills or general ability. These results match other accounts of positive selection into stand alone online courses in observational studies. Our preferred estimates suggest the effects are around 0.4 points translating into two-thirds of a standard deviation of average GPA in the sample.

In the second row, we look at GPA conditional on graduation, i.e., GPA among graduates only. We recognize graduation is endogenous and in fact find a significant negative impact of distance education on graduation. But, for completeness we report the GPA results conditional on graduation. The effects of distance education are smaller in magnitude among the sample of graduates suggesting a reduction of 0.08 points in average GPA in our preferred specification (column (4)). This translates into just over 1/10th of a standard deviation of average GPA. In terms of graduation, distance students are 15 percentage points less likely to graduate. They are also more likely to request thesis extensions conditional on graduation.

Broadly, these results point to a large and negative impact of distance education in academic programs that are consistent across outcomes namely average GPA, graduation, late graduation and number of thesis extensions. Moreover, the negative effects of distance education continue to impact military officers into their career. Officers enrolled in distance programs at NPS are 12 percentage points less likely to promote and are 9 percentage points more likely to separate. Since the military regularly sends officers to NPS for graduate education, they are familiar with the degree programs offered at the school. In these cases the degrees are less likely to represent a signal to future employers as they would in other labor markets. Unlike recent studies that have documented a negative effect of online degrees on job call-backs, our labor market results suggest students in distance programs have worse learning outcomes (as evidenced by their average GPA) and performed worse in their subsequent military jobs. Thus, the findings across academic and labor market outcomes are consistent and suggest productivity impacts of distance education.

5.3 Treatment Heterogeneity

In Table 3, we reported average treatment effects of distance education. However, there is no reason to believe the impact of distance education is similar across groups. To test for heterogeneity in the treatment effects, we split the sample into various sub-samples based on demographics, academic preparation, military service and community, and military rank. Within each sub-sample, we estimate the same nearest neighbor matching regressions with Abadie-Imbens robust standard errors. The coefficients on distance education in these sub-samples are reported in Tables 4 to 7. As noted earlier we do not have sufficient overlap for some sub-samples by race and community. We report the complete results in the tables, but in the discussion below we focus on results where we do observe significant overlap.

Table 4 focuses on treatment heterogeneity by demographics. Although the coefficients on distance education for the academic outcomes in the female sample are larger, the estimates are not as precisely estimated. The 95 percent confidence interval for average GPA includes a positive effect of 0.07 points on the high end to a negative effect of 1 point on the low end. Moreover, we find no significant impacts of distance education on labor market outcomes by demographic group. Interestingly, we find smaller negative effects of distance education and insignificant effects on job outcomes for Hispanics. Overall, the negative impacts of distance programs are less severe for Hispanic students compared to white students. This finding is different from other settings where non-white students tend to perform worse in online courses. The coefficients on married are similar to those in Table 3 indicating no significant heterogeneous effects by marital status even though this group is more likely to attend distance programs.

In Table 5 we report the results by academic preparation namely the type of undergraduate institution and assigned NPS Academic Profile Codes. Again we focus on results where we have sufficient power in the data (namely numbers of distance education students) and overlap in the matching exercise. Students who earned their degrees from private non-profit universities and service academies are slightly better off in distance programs in terms of average GPA compared to students from public universities. But, these results are not robust across the different outcomes. For example, the results on promotion indicate remarkably similar negative effects of distance programs ranging from 11 to 14 percentage points. We find heterogeneous effects by prior academic preparation, but again it is unclear how to interpret these effects because they are not consistent across the set of outcomes.

We find more consistent results by military community in Table 6. As we discussed earlier, military pilots are positively selected based on their initial ability and technical training. In this sample of higher ability officers, we find small negative impacts of distance education on average GPA and promotion. In comparison, we find very large and negative effects of distance programs on submarine and support officers across the board, from academic to job-related outcomes. Finally, we look at heterogeneous effects by officer rank in Table 7. As we move up officer ranks, the negative impact of distance education on average GPA decreases from 0.9 points for O1s to 0.24 points for O5s. We observe a similar pattern by rank on promotion as seen in column (5). Broadly, these results point to significant heterogeneity by group. The negative impacts of distance education are more severe for some military students, for example, submarine officers, those that attended public undergraduate institutions, and junior officers. In comparison the effects are relatively small and insignificant for Hispanics and more senior officers.

6 Conclusion

Using a unique student-level dataset on academic and job-related outcomes, we estimate the impact of distance education programs on economic outcomes ranging from graduation to job-promotion. In such studies the main challenge of estimating causal effects is the non-random assignment of students to different program types. To address this problem, we implement propensity score matching using a rich set of observable characteristics of students. Our matching exercise suggests a large, negative and significant impact of distance education programs on the academic and subsequent job performance of military officers. Students that enroll in distance programs at NPS have lower average GPAs, are less likely to graduate, more likely to graduate late conditional on graduation, are less likely to promote in their military career and are more likely to separate from the military. But, the negative impacts of distance education vary widely by group within the military and by prior academic preparation of students. Although our results thus far are interesting, we hope to identify specific factors contributing to these negative effects by adding more data on courses, degree programs and the range of delivery modes within the distance programs.

7 References

1. Abadie, A. & Imbens, G. (2006). Large sample properties of matching estimators for average treatment effects. Econometrica, 74(1), 235-267.

2. Abadie, A. & Imbens, G. (2009). Matching on the estimated propensity score. NBER Working Paper 15301. Revised 2011.

3. Alpert, W. T., Couch, K. A., & Harmon, O. R. (2015). A randomized assessment of online learning. Storrs, CT: University of Connecticut.

4. Bernard, R. M., Abrami, P. C., Lou, Y., Borokhovski, E., Wade, A., Wozney, L., Wallet, P. A., Fiset, M., & Huang, B. (2004). How does distance education compare with classroom instruction? A meta-analysis of the empirical literature. Review of Educational Research, 74 (3), 379–439.

5. Brown, B. W. & Liedholm, C. E. (2002). Teaching microeconomic principles: can web courses replace the classroom in Principles or Microeconomics? American Economic Review, 92(2), 444-448.

6. Bowen, W. G., Chingos, M. M., Lack, K. A., & Nygren, T. I. (2014). Interactive learning online at public universities: Evidence from a six-campus randomized trial. Journal of Policy Analysis and Management, 33(1), 94-111.

7. Coates, D., Humphreys, B. R., Kane, J., & Vachris, M. A. (2004). "No significan distance" between face-to-face and online instruction: evidence from principles of economics. Economics of Education Review, 23, 533-546.

8. Cosgrove, Sarah and Neal H. Olitsky (2015). "Knowledge Retention, Student Learning, and Blended Course Work: Evidence from Principles of Economics Courses." Southern Economic Journal, 82(2), 556-579.

9. Figlio, D. N., Rush, M., & Yin, L. (2010). Is it live or is it internet? experimental estimates of the effects of online instruction on student learning. Journal of Labor Economics, 31(4), 763-784.

10. Gratton-Lavoie, C. & Stanley, D. (2009). Teachign and learning principles of microeconomics online: an empirical assessment. Journal of Economic Education, 40(1), 3-25.

11. Harmon, O. R., Alpert, W. T., & Lambrinos, J. (2014). Testing the effect of hybrid lecture delivery on learning outcomes. Journal of Online Learning and Teaching, 10(1), 112-121.

12. Joyce, Theodore. J., Sean Crockett, S., Jaeger, D. A., Altindag, O., & O'Connell, S. D. (2014). Does classroom time matter? A randomized field experiment of hybrid and traditional lecture formats in economics. National Bureau of Economic Research, Inc, NBER Working Paper 20006.

13. Koch, J. V. (2006). Public investment in university distance learning programs: some performance-based evidence. Atlantic Economic Journal, 34, 23-32.

14. Lack, K. A. (2013). Current status of research on online learning in postsecondary education. New York, NY: Ithaka S+R.

15. Means, B., Toyama, Y., Murphy, R., Bakia, M., & Jones, K. (2010). Evaluations of evidence-based practices in online learning: a meta-analysis and review of online learning studies. Washington, D.C.: U.S. Department of Education.

16. Olitsky, Neal. H. & Cosgrove, Sarah. B. (2015). The effect of blended courses on student learning: evidence from introductory economic courses. Internation(al Review of Economics Education, 15, 17-31.

8 Figures

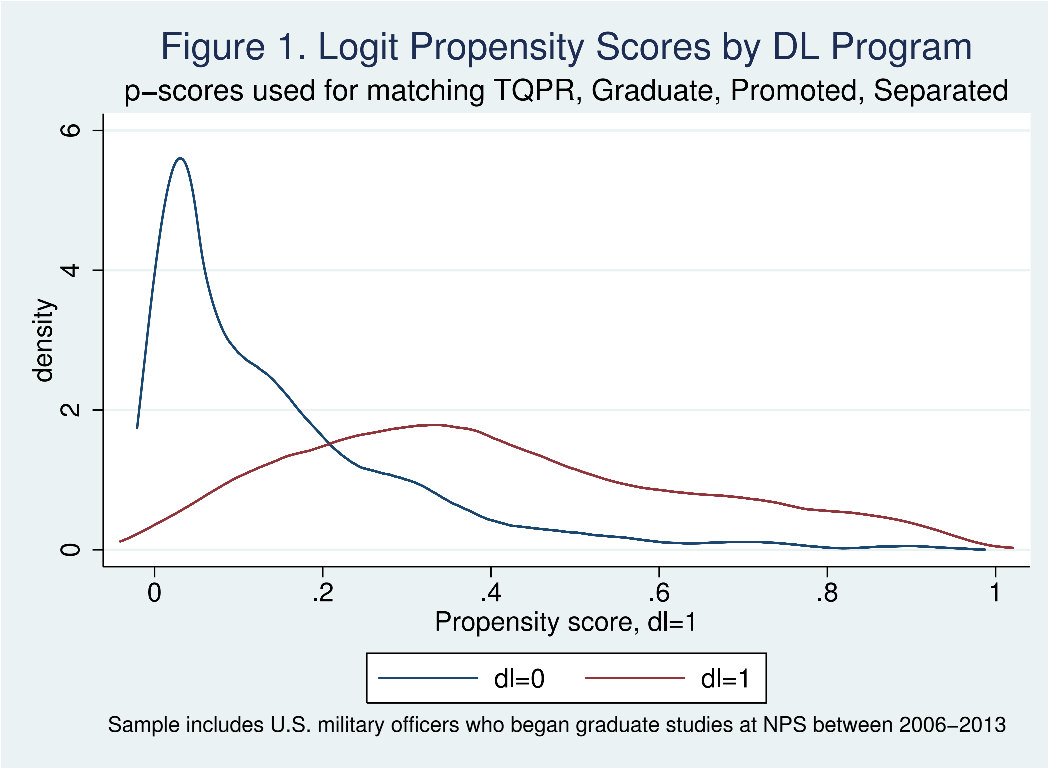
Figure 1. Logit Propensity Scores by DL Program
p−scores used for matching TQPR, Graduate, Promoted, Separated
density
0
2
4
6
0
.2
.4
.6
.8
1
Propensity score, dl=1
dl=0
dl=1
Sample includes U.S. military officers who began graduate studies at NPS between 2006−2013

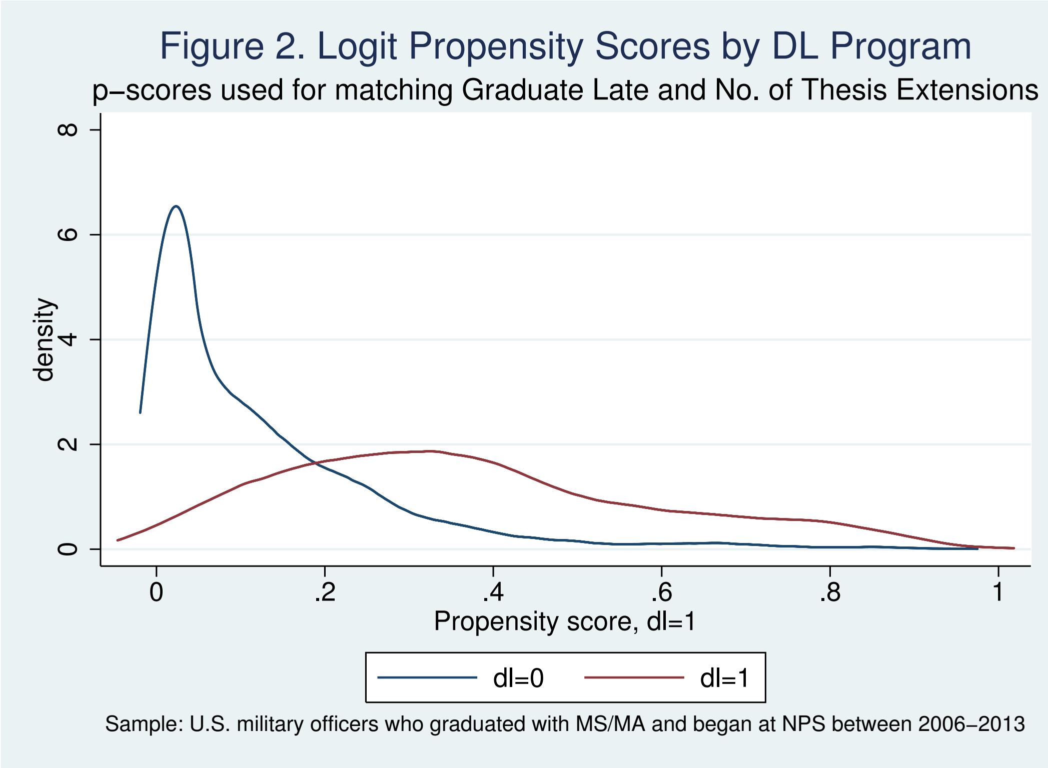
Figure 2. Logit Propensity Scores by DL Program
p−scores used for matching Graduate Late and No. of Thesis Extensions
density
0
2
4
6
8
0
.2
.4
.6
.8
1
Propensity score, dl=1
dl=0
dl=1
Sample: U.S. military officers who graduated with MS/MA and began at NPS between 2006−2013

Figure 3A. Overlap Graphs for Matching by Demographics

p−scores used for matching TQPR, Graduate, Promoted, Separated

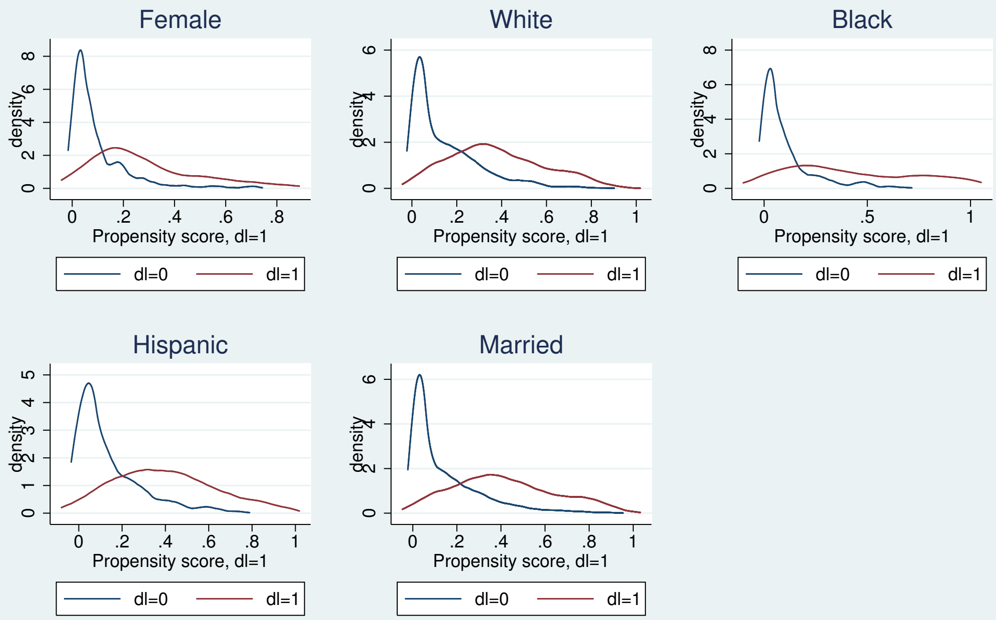

Figure 3B. Overlap Graphs for Matching by Demographics

p−scores used for matching Graduate Late and No of Thesis Extensions

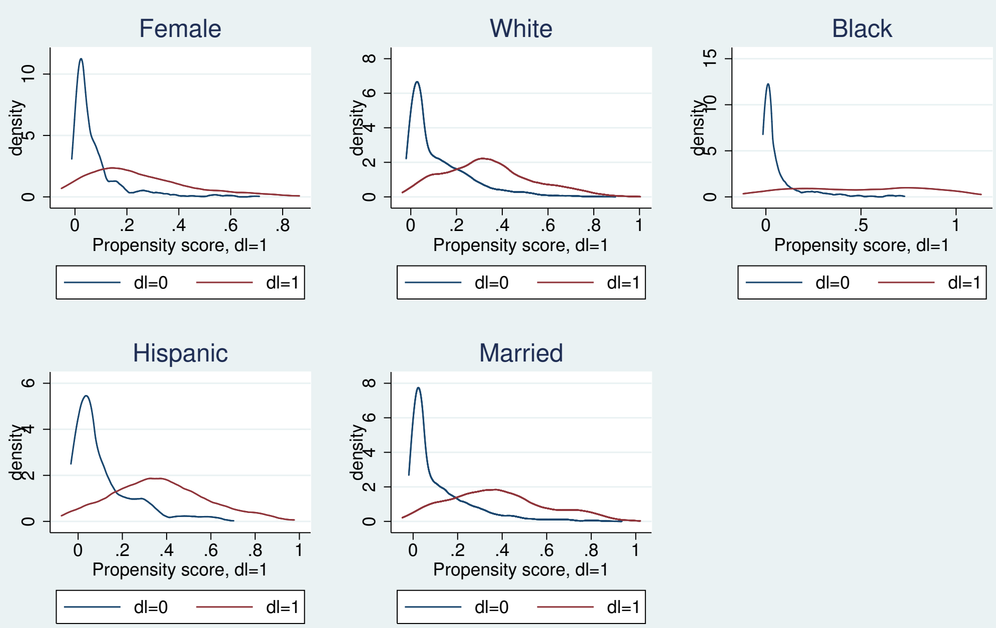

Figure 4A. Overlap Graphs for Matching by Academic Prep

p–scores used for matching TQPR, Graduate, Promoted, Separated

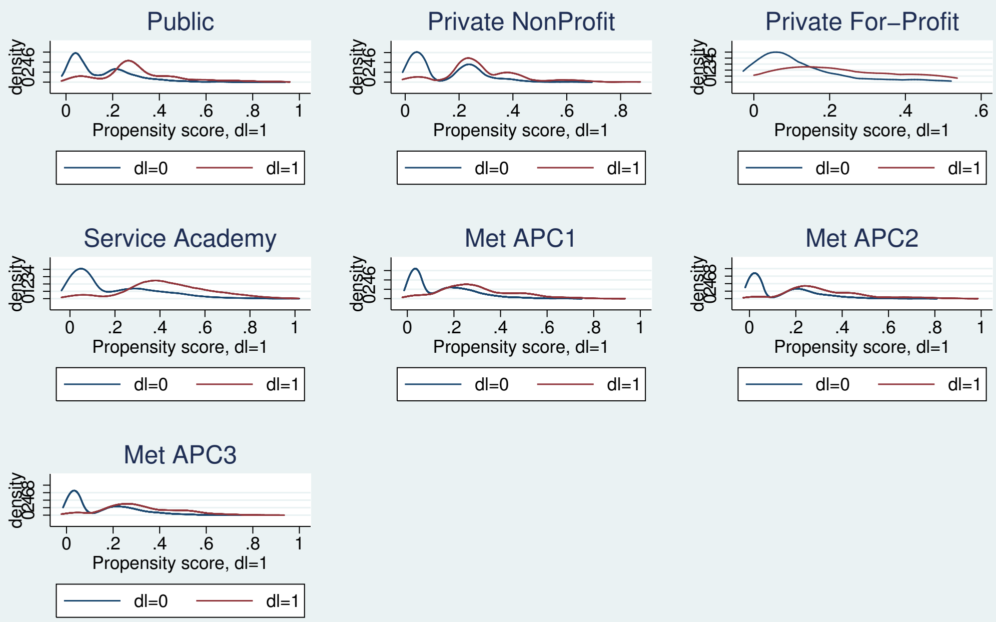

Figure 4B. Overlap Graphs for Matching by Academic Prep

p–scores used for matching Graduate Late and No of Thesis Extensions

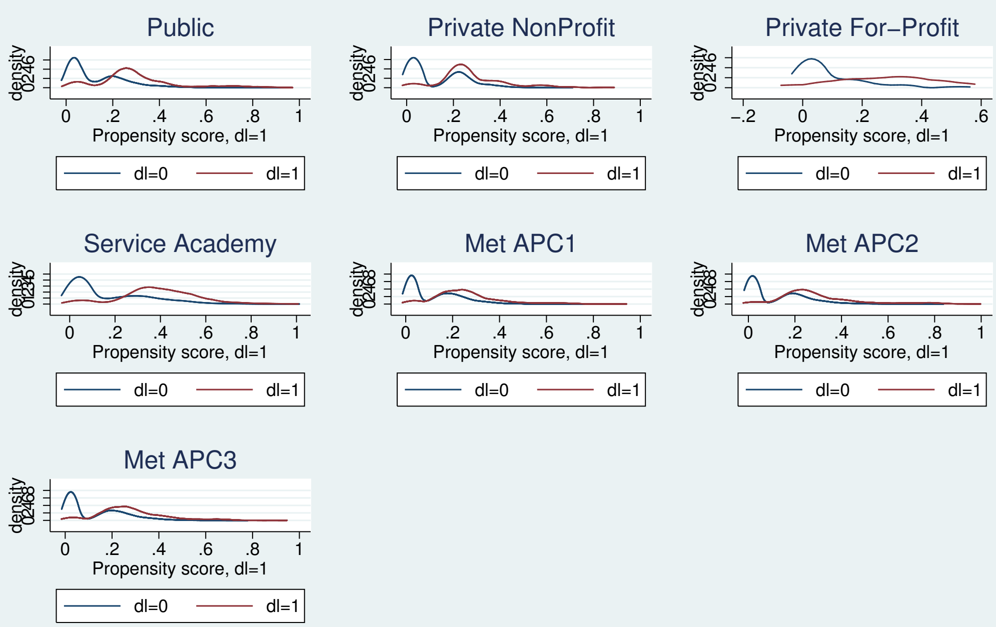

Figure 5A. Overlap Graphs for Matching by Service

p–scores used for matching TQPR, Graduate, Promoted, Separated

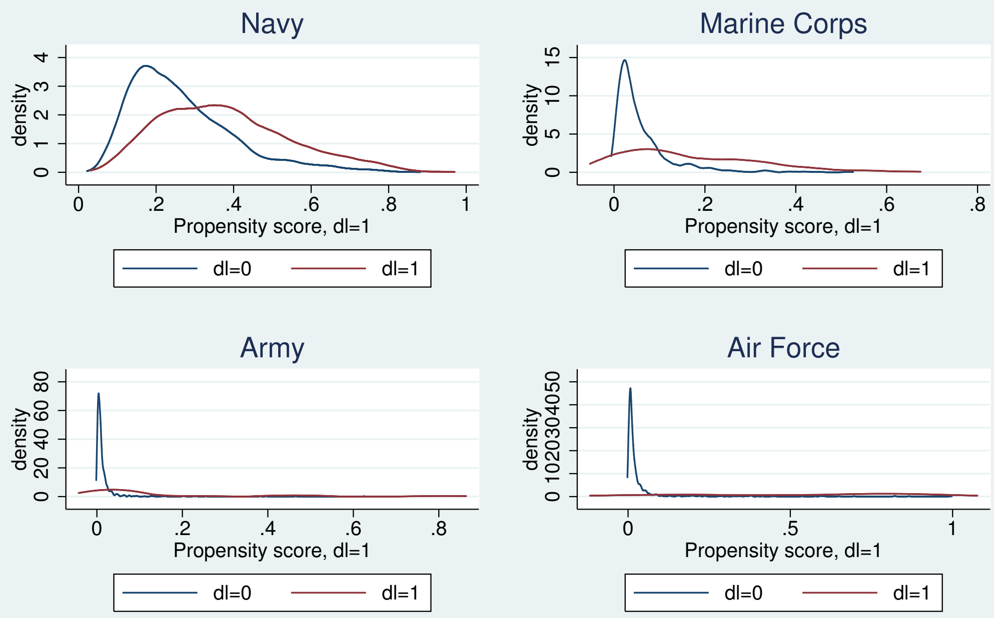

Figure 5B. Overlap Graphs for Matching by Service

p–scores used for matching Graduate Late and No of Thesis Extensions

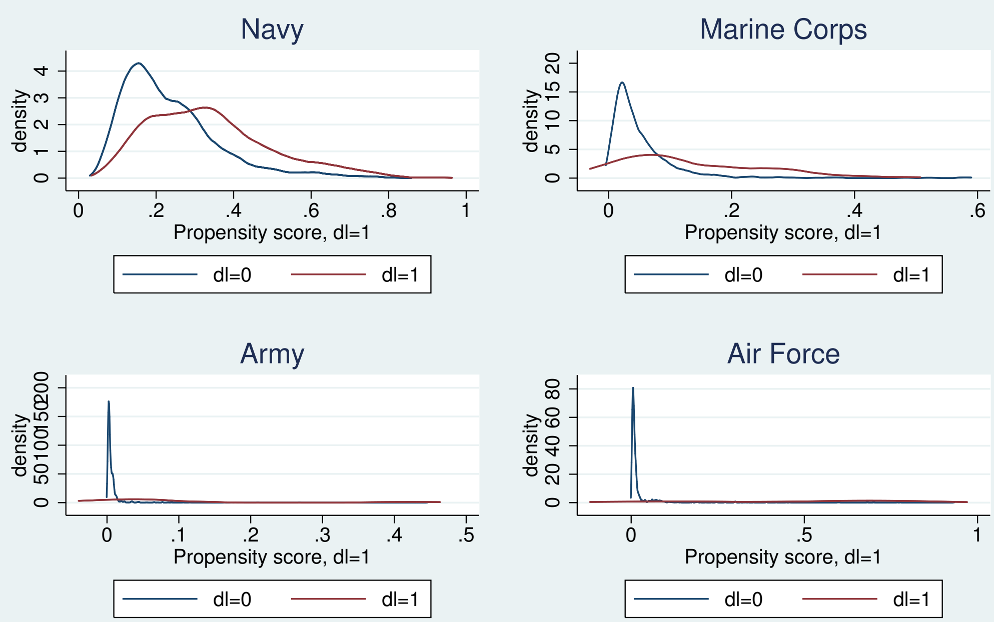

Figure 6A. Overlap Graphs for Matching by Community

p–scores used for matching TQPR, Graduate, Promoted, Separated

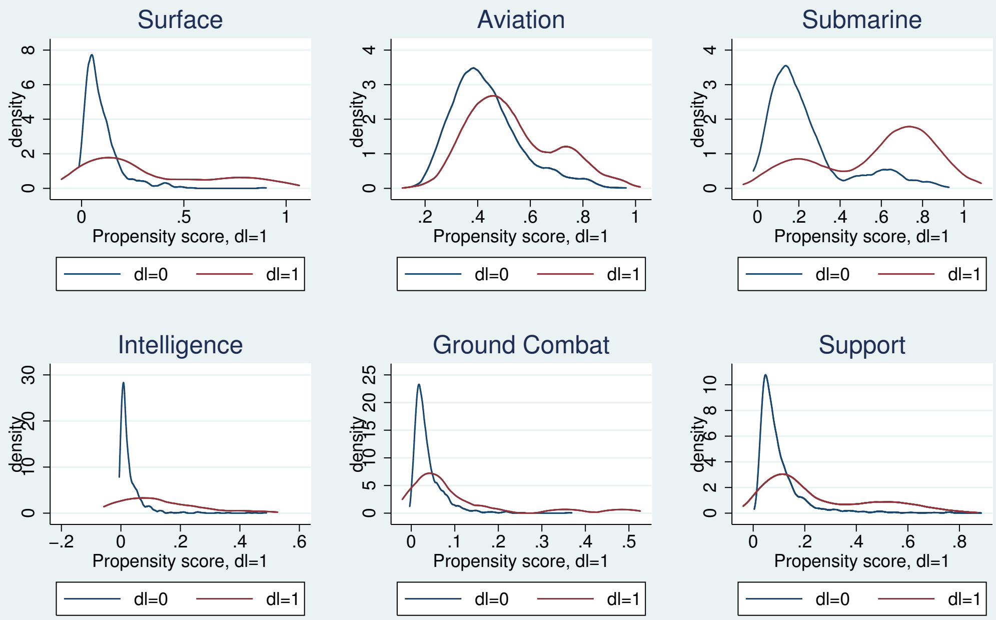

Figure 6B. Overlap Graphs for Matching by Community

p–scores used for matching Graduate Late and No of Thesis Extensions

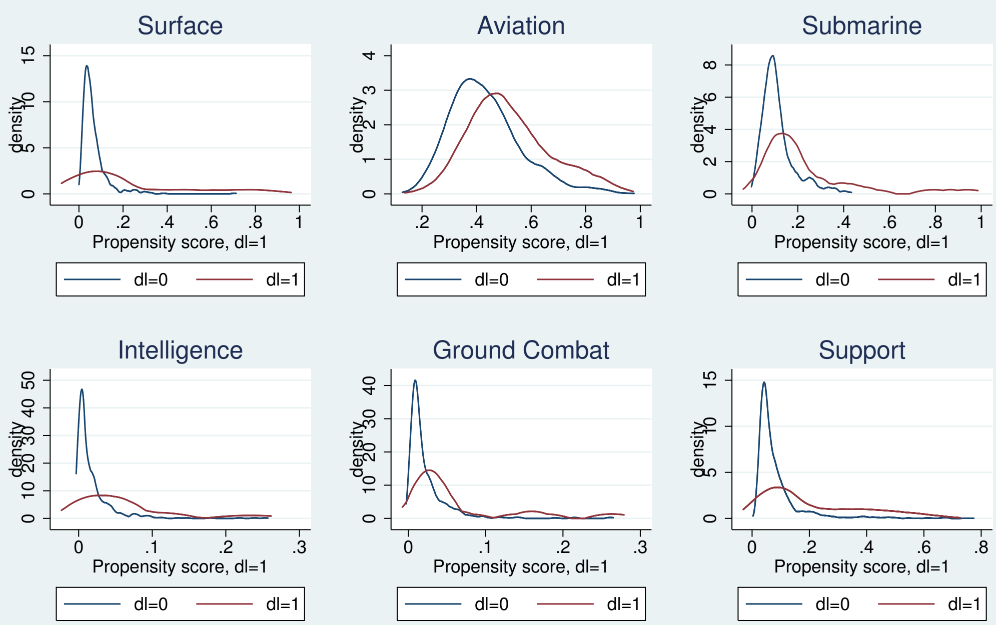

Figure 7A. Overlap Graphs for Matching by Officer Rank

p–scores used for matching TQPR, Graduate, Promoted, Separated

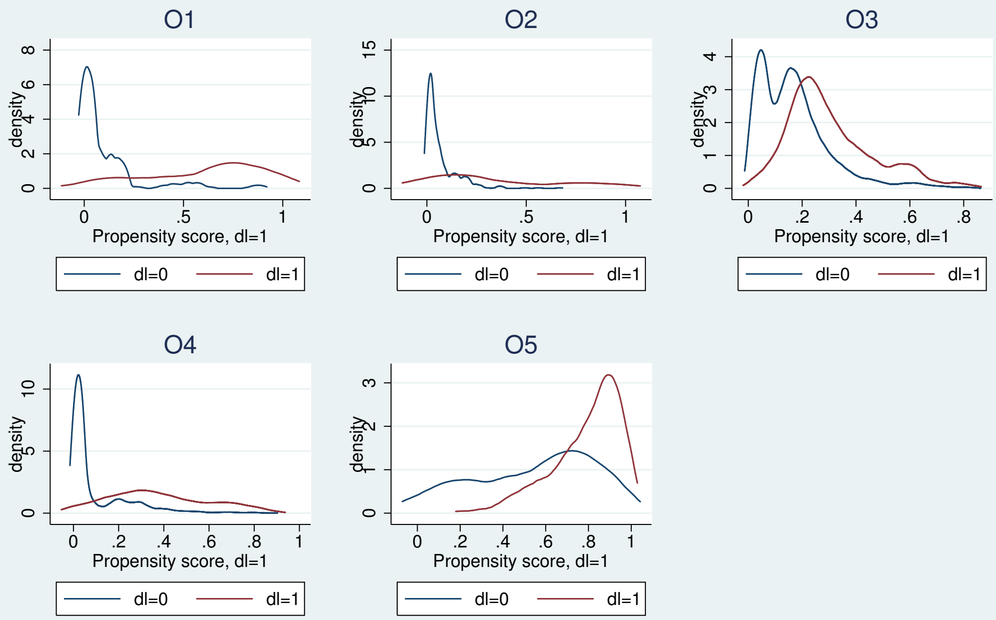

Figure 7B. Overlap Graphs for Matching by Officer Rank

p–scores used for matching Graduate Late and No of Thesis Extensions

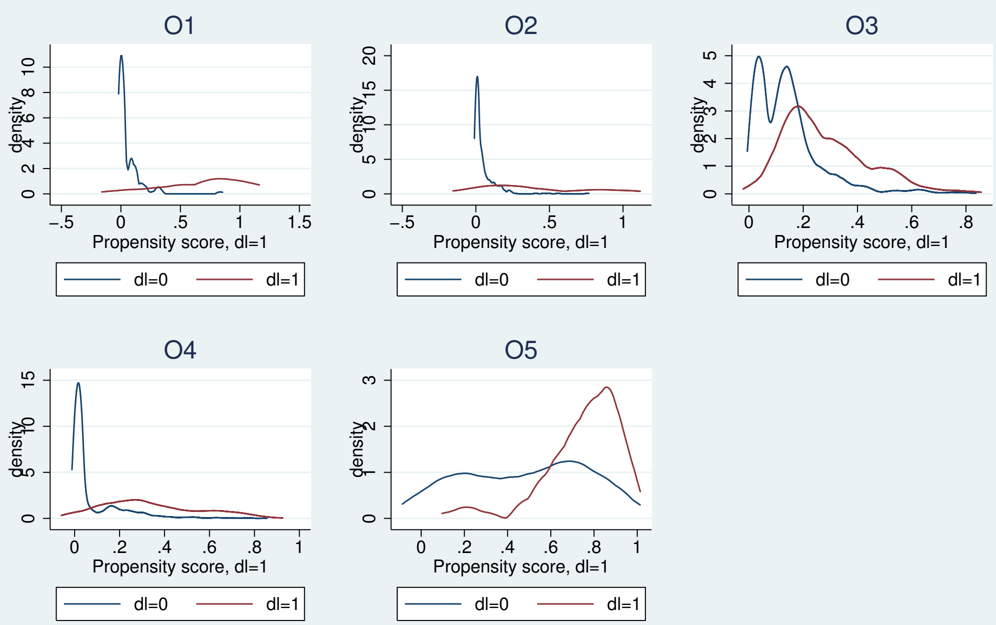

Table 1: Descriptive Statistics

	Residence		Distance		All	
	Mean	SD	Mean	SD	Mean	SD
Outcomes						
GPA	3.693	0.3	3.402	1.052	3.636	0.551
Graduate	0.851	0.356	0.686	0.464	0.818	0.386
Graduate Late	0.044	0.205	0.046	0.21	0.044	0.205
# Thesis Extensions	0.065	0.334	0.103	0.5	0.071	0.367
Promoted	0.436	0.496	0.292	0.455	0.408	0.491
Separated	0.084	0.277	0.101	0.302	0.087	0.282
Demographics						
Female	0.096	0.295	0.053	0.223	0.088	0.283
White	0.741	0.438	0.713	0.453	0.736	0.441
Black	0.066	0.248	0.048	0.214	0.062	0.242
Hispanic	0.069	0.254	0.062	0.241	0.068	0.251
OtherRace	0.124	0.329	0.177	0.382	0.134	0.341
Married	0.731	0.444	0.722	0.448	0.729	0.444
Divorced	0.019	0.136	0.003	0.055	0.016	0.124
Age	31.829	6.783	29.972	11.272	31.463	7.906
Missing Age	0.024	0.154	0.1	0.3	0.039	0.194
# Dependents	1.787	1.547	1.688	1.556	1.767	1.549
Missing Dependents	0.048	0.214	0.138	0.345	0.066	0.248
Years since UG	6.934	4.286	7.784	5.593	7.102	4.588
Missing Yrs UG	0.045	0.207	0.121	0.326	0.06	0.237
Job-Related						
Navy	0.511	0.5	0.888	0.315	0.586	0.493
MarineCorps	0.141	0.349	0.041	0.197	0.122	0.327
Army	0.168	0.374	0.012	0.109	0.137	0.344
AirForce	0.164	0.37	0.052	0.222	0.142	0.349
Support	0.452	0.498	0.253	0.435	0.412	0.492
Surface	0.131	0.337	0.078	0.268	0.121	0.326
Submarine	0.041	0.199	0.092	0.289	0.051	0.221
Aviation	0.141	0.348	0.524	0.5	0.216	0.412
Intelligence	0.127	0.333	0.018	0.133	0.106	0.307
Ground Combat	0.143	0.35	0.025	0.156	0.12	0.325

Note: UG refers to undergraduate degree. The summary statistics for Graduate Late and # Thesis Extensions are conditional on graduation.

Table 1: Descriptive Statistics (cont'd)

	Residence		Distance		All	
Variable	Mean	SD	Mean	SD	Mean	SD
Job-Related Cont'd						
O1	0.022	0.148	0.022	0.146	0.022	0.148
O2	0.087	0.282	0.035	0.183	0.077	0.266
O3	0.532	0.499	0.53	0.499	0.531	0.499
O4	0.318	0.466	0.204	0.403	0.296	0.456
O5	0.009	0.096	0.098	0.297	0.027	0.161
O6	0.001	0.027	0.011	0.102	0.003	0.052
Academic						
Met APC1 requirement	0.897	0.304	0.766	0.424	0.871	0.335
Met APC2 requirement	0.768	0.422	0.635	0.482	0.742	0.438
Met APC3 requirement	0.901	0.298	0.814	0.389	0.884	0.32
Undergraduate Institution:						
Service Academy	0.227	0.419	0.31	0.462	0.243	0.429
Public	0.629	0.483	0.554	0.497	0.614	0.487
Private NonProfit	0.205	0.404	0.171	0.377	0.198	0.399
Private ForProfit	0.011	0.106	0.007	0.082	0.011	0.102
Observations	5,423		1,331		6,754	
Matched Obs.	5,353		1,324		6,678	

Note: APC refers to Academic Proficiency Codes assigned by NPS registrar's office. See text for more details on APC and other data sources.

Table 2: Logit Estimates from First Stage Matching
Dep Var: Indicator for Distance Education

	(1) All	(2) Graduated
Female	-0.4240***	-0.5658***
	[0.1513]	[0.1849]
Black	-0.2365	-0.2724
	[0.1540]	[0.1803]
Hispanic	-0.0053	-0.0763
	[0.1429]	[0.1720]
OtherRace	-0.1526	-0.2046
	[0.1300]	[0.1610]
Married	0.4469***	0.4371***
	[0.1149]	[0.1350]
Divorced	-0.0315	-0.3564
	[0.5313]	[0.7151]
Age	0.0172	0.0125
	[0.0124]	[0.0147]
xAGE	1.0460**	0.9282
	[0.5298]	[0.6080]
Dependents	-0.0489	-0.063
	[0.0326]	[0.0391]
xDEPS	1.2855***	1.2611***
	[0.2431]	[0.2613]
YrsFrUGrad	0.0566***	0.0623***
	[0.0107]	[0.0122]
xUGradYr	0.9920***	0.4892**
	[0.1556]	[0.2034]
Public	-0.4718***	-0.4994***
	[0.0836]	[0.0970]
ServiceAcademy	0.6291***	0.6793***
	[0.0877]	[0.1029]
MarineCorps	-1.4992***	-1.2493***
	[0.1820]	[0.1982]
Army	-3.7506***	-4.1501***
	[0.3521]	[0.5080]
AirForce	-2.6506***	-2.4362***
	[0.3401]	[0.3413]

Robust standard errors reported in brackets. *** p<0.01, ** p<0.05, * p<0.1

Table 2: Logit Estimates from First Stage Matching (cont'd)
Dep Var: Indicator for Distance Education

	(1) All	(2) Graduated
Support	-0.8996	-1.5044*
	[0.6009]	[0.7968]
Navy Support	0.0649	0.7071
	[0.6028]	[0.7987]
Marine Support	0.146	0.0844
	[0.6712]	[0.8933]
Army Support	1.2447	1.6184
	[0.7809]	[1.1003]
AirForce Support	2.2129***	2.2712***
	[0.7007]	[0.8794]
O1	0.3553	0.0463
	[0.2724]	[0.3309]
O2	-0.4863***	-0.4635**
	[0.1736]	[0.2058]
O4	-0.0814	-0.0946
	[0.1086]	[0.1270]
O5	1.8487***	1.6261***
	[0.2379]	[0.2588]
O6	1.3725*	1.1116
	[0.7050]	[0.7453]
yr2007	0.5103***	0.5120***
	[0.1712]	[0.1821]
yr2008	0.5507***	0.4958***
	[0.1724]	[0.1824]
yr2009	0.8094***	0.7370***
	[0.1680]	[0.1794]
yr2010	0.2319	0.1443
	[0.1762]	[0.1916]
yr2011	1.1555***	0.9876***
	[0.1614]	[0.1753]
yr2012	0.9722***	0.8583***
	[0.1632]	[0.1807]
yr2013	0.6509***	0.3841
	[0.1665]	[0.2550]
Constant	-2.5801***	-2.4297***
	[0.3981]	[0.4620]
Observations	6,678	5,457

Robust standard errors reported in brackets. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 3: OLS and Propensity Score Matching Estimates
Dependent Variable: Indicator for Distance Education

	(1)	(2)	(3)	(4)
	Standard OLS	Inverse P-Score Weights	Inverse P-Score Weights	Nearest Neighbor Matching, Abadie-Imbens Std Errors
GPA	-0.2722***	-0.4292***	-0.5013***	-0.3855***
	[0.0292]	[0.0833]	[0.0647]	[0.1322]
GPA, condt'l on G	0.0264***	-0.0170	-0.0487***	-0.0785***
	[0.0079]	[0.0161]	[0.0168]	[0.0302]
Graduate=1	-0.1260***	-0.1667***	-0.2293***	-0.1508***
	[0.0134]	[0.0237]	[0.0236]	[0.0432]
Graduate Late	0.0056	0.0492***	0.0780***	0.0825**
	[0.0095]	[0.0159]	[0.0200]	[0.0410]
# Thesis Extensions (condt'l)	0.0426*	0.1261***	0.1840***	0.2289**
	[0.0218]	[0.0348]	[0.0464]	[0.0889]
Promoted	-0.0162	-0.0767***	-0.1312***	-0.1158***
	[0.0144]	[0.0195]	[0.0202]	[0.0191]
Separated	0.0208**	0.0526***	0.0157	0.0902**
	[0.0104]	[0.0183]	[0.0172]	[0.0409]
Sample	Matched	Unmatched	Matched	Unmatched
No of Obs	6,613	6,678	6,613	6,678
No of Obs, condtl on G	5,400	5,457	5,457	5,457

Notes: *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Standard errors are reported in brackets, with column (4) accounting for first-stage estimation of p-scores. All regressions include the following regressors: female, Black, Hispanic, OtherRace, Married, Divorced, Age, No of Dependents, Years since received Undergrad Degree, Undergrad Institution was Public, Service Academy graduate, Marine Corps, Army, Air Force, interactions of service branch and Support community, dummies for Officer Rank, indicators for having met APC1, APC2 and APC3 qualifications, cohort fixed effects, and indicators for missing variables.

Table 4: Heterogeneity in Treatment Effects by Demographic Characteristics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	GPA	Graduation	Late Graduation	# Thesis Extensions	Promoted	Separated	Observations
Female	-0.4881*	-0.1201***	0.3159***	0.5293**	-0.1887	-0.0257	583
	[0.2837]	[0.0254]	[0.0487]	[0.2218]	[0.1656]	[0.0288]	
White	-0.3854***	-0.1476***	0.1742***	0.3942***	-0.1144***	0.0478	4,922
	[0.1142]	[0.0379]	[0.0568]	[0.1504]	[0.0280]	[0.0302]	
Black	-0.4045***	-0.3756***	-0.0781	-0.1081*	-0.1794***	0.0120	418
	[0.0844]	[0.0535]	[0.0502]	[0.0584]	[0.0335]	[0.0427]	
Hispanic	-0.1684**	-0.2477***	-0.0342	0.0085	-0.0886	-0.0477**	440
	[0.0773]	[0.0711]	[0.0388]	[0.1098]	[0.0901]	[0.0240]	
Married	-0.3273***	-0.1721***	0.1723***	0.4467***	-0.1333***	0.0484	4,873
	[0.0931]	[0.0420]	[0.0504]	[0.1419]	[0.0246]	[0.0303]	

Notes: *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Standard errors are reported in brackets, accounting for first-stage estimation of p-scores per Abadie and Imbens (2015). All regressions control for years since received undergrad degree, public undergrad institution, Service Academy graduate, branch of service, community, service and community interactions, rank, APC qualifications, cohort fixed effects, and indicators for missing variables. The regressions in columns (3) and (4) are conditional on graduation.

Table 5: Heterogeneity in Treatment Effects by Academic Preparation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	GPA	Graduation	Late Graduation	# Thesis Extensions	Promoted	Separated	Observations
Public	-0.3271***	-0.1648***	0.1149***	0.3858***	-0.1067***	0.0444*	4,120
	[0.0688]	[0.0364]	[0.0298]	[0.0896]	[0.0345]	[0.0243]	
Private NonProfit	-0.2764***	-0.1577***	0.1913***	0.4235**	-0.0821	0.0416	1,327
	[0.1038]	[0.0457]	[0.0546]	[0.2129]	[0.0659]	[0.0366]	
Private For-Profit	-0.1100***	-0.3913*	0.7091	2.0364	-0.2319**	-0.0435*	69
	[0.0290]	[0.2349]	[0.4633]	[1.3818]	[0.0952]	[0.0244]	
Service Academy	-0.2692***	-0.1048***	0.1122***	0.4913***	-0.1443***	0.0788*	1,629
	[0.0902]	[0.0405]	[0.0306]	[0.1546]	[0.0519]	[0.0415]	
On Admission to NPS, Academic Profile Code:							
Met APC1	-0.3483***	-0.1753***	0.1082**	0.3323**	-0.1501***	0.0678*	5,818
	[0.0844]	[0.0305]	[0.0487]	[0.1403]	[0.0330]	[0.0368]	
Met APC2	-0.4953***	-0.1701***	0.0377	0.1545***	-0.1213***	0.0320	4,948
	[0.1171]	[0.0344]	[0.0281]	[0.0560]	[0.0449]	[0.0238]	
Met APC3	-0.4296***	-0.1523***	0.0761	0.1915*	-0.1225***	0.0461	5,901
	[0.0995]	[0.0313]	[0.0591]	[0.0992]	[0.0384]	[0.0318]	
NO APC1	-0.2224*	-0.1558***	0.1722***	0.3592***	-0.1308***	0.0791**	860
	[0.1209]	[0.0449]	[0.0481]	[0.0967]	[0.0402]	[0.0333]	
NO APC2	-0.2497***	-0.1876***	0.1208***	0.2831***	-0.0623*	0.0194	1,730
	[0.0470]	[0.0376]	[0.0232]	[0.0948]	[0.0366]	[0.0218]	
NO APC3	-0.2144***	-0.1396***	0.1524***	0.3414***	-0.0830*	0.0412	777
	[0.0773]	[0.0493]	[0.0348]	[0.0778]	[0.0434]	[0.0316]	

Notes: *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Standard errors are reported in brackets, accounting for first-stage estimation of p-scores per Abadie and Imbens (2015). All regressions control for demographic characteristics, service, community, service and community interactions, rank, cohort fixed effects, and indicators for missing variables. The regressions in columns (3) and (4) are conditional on graduation.

Table 6: Heterogeneity in Treatment Effects by Service and Community

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	GPA	Graduation	Late Graduation	# Thesis Extensions	Promoted	Separated	Observations
Navy	-0.2675***	-0.1229***	-0.0392***	-0.0408**	-0.0212	0.0132	3,923
	[0.0406]	[0.0203]	[0.0077]	[0.0204]	[0.0224]	[0.0125]	
MarineCorps	-0.3897*	-0.1147	0.0688	0.0578	-0.1409*	0.0923	802
	[0.2113]	[0.0893]	[0.1671]	[0.1678]	[0.0799]	[0.0748]	
Army	-0.9293	-0.6727***	-0.0415***	-0.0591***	-0.3942***	0.0351	912
	[1.3782]	[0.0394]	[0.0072]	[0.0111]	[0.0174]	[0.0662]	
AirForce	-0.3067*	-0.0487	0.3237	0.8491	-0.2148	0.1683	945
	[0.1574]	[0.0651]	[0.2848]	[0.8149]	[0.2728]	[0.2559]	
Surface	-0.5923***	-0.2460***	-0.0665***	-0.0870***	-0.1924***	0.0197	811
	[0.2061]	[0.0475]	[0.0103]	[0.0155]	[0.0445]	[0.0379]	
Submarine	-1.1039***	-0.4186***	-0.0599***	-0.0876***	-0.2471***	-0.0465**	344
	[0.1581]	[0.0584]	[0.0184]	[0.0322]	[0.0569]	[0.0208]	
Aviation	-0.0761***	0.0170	-0.0170	-0.0331	-0.0520*	0.0476**	1,452
	[0.0253]	[0.0238]	[0.0132]	[0.0213]	[0.0279]	[0.0214]	
Intelligence	-1.0560*	-0.2940*	-0.0241	-0.0328	-0.1392***	0.0170	704
	[0.5514]	[0.1777]	[0.0966]	[0.0968]	[0.0199]	[0.0708]	
GroundCombat	-1.3185***	-0.4102***	-0.0119	-0.0015	-0.1633	0.0138	796
	[0.4575]	[0.1323]	[0.0206]	[0.0474]	[0.1252]	[0.0356]	
Support	-0.3010***	-0.1746***	0.0673*	0.2131**	-0.1557***	0.0600*	2,743
	[0.0980]	[0.0408]	[0.0405]	[0.0981]	[0.0375]	[0.0317]	

Notes: *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Standard errors are reported in brackets, accounting for first-stage estimation of p-scores per Abadie and Imbens (2015). All regressions control for demographic characteristics, years since undergrad degree, public undergrad institution, Service Academy graduate, APC qualifications, cohort fixed effects, and indicators for missing variables. The regressions in columns (3) and (4) are conditional on graduation.

Table 7: Heterogeneity in Treatment Effects by Military Officer Rank

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	GPA	Graduation	Late Graduation	# Thesis Extensions	Promoted	Separated	Observations
O1	-0.8717***	-0.3510**	0.1654	0.3534	-0.3510**	0.0000	151
	[0.1700]	[0.1644]	[0.4173]	[0.8343]	[0.1676]	[0.0126]	
O2	-0.4060***	-0.1696	0.0602	0.2477	-0.1024*	0.0690	519
	[0.1189]	[0.1515]	[0.0872]	[0.2629]	[0.0620]	[0.0454]	
O3	-0.2173***	-0.1164***	0.0144	0.0647	-0.0873**	0.0704***	3,589
	[0.0439]	[0.0269]	[0.0275]	[0.0446]	[0.0345]	[0.0260]	
O4	-0.3841	-0.1672**	0.1258	0.2998**	-0.1352**	0.0108	1,998
	[0.2598]	[0.0772]	[0.0776]	[0.1198]	[0.0651]	[0.0304]	
O5	-0.2413*	-0.2556***	-0.0317	-0.1111	0.0333	-0.1056*	180
	[0.1360]	[0.0542]	[0.0544]	[0.1389]	[0.0312]	[0.0619]	

Notes: *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Standard errors are reported in brackets, accounting for first-stage estimation of p-scores per Abadie and Imbens (2015). All regressions control for demographic characteristics, years since undergrad degree, public undergrad institution, Service Academy graduate, APC qualifications, cohort fixed effects, and indicators for missing variables. The regressions in columns (3) and (4) are conditional on graduation.